コンジョイント分析における効用関数について (不確実性の下での数理モデルの構築と最適化)

上田, 徹

数理解析研究所講究録
1194巻 2001年 184-192

コンジョイント分析における効用関数について

成蹊大学　上田　徹 (Tohru Ueda)
Seikei University

1. まえがき

コンジョイント分析法は古くから心理学の分野で議論されてきたが、最初にLuce, Tukey [1] がコンジョイント測定法と呼んだ「ある与えられた結合ルールのもとで、部分効用が数値的に表現されるためには、与えられた順序関係がどのような性質を満たしていなければならないか」を考える公理論的体系化の試みから発展を促され、Kruskalの単調回帰原理を用いた実用的アルゴリズムの提案 [2] 以来、マーケティングの分野でも幅広く使われるようになった。

コンジョイント分析では選好順位と効用関数の値の間に単調関係が成立するよう効用関数のパラメタを調整する。しかし、単調関係が容易に見いだせないような選好順位データも存在する。そのような場合には単調関係のずれを評価する尺度を設け、一定の範囲に収まっていれば許容するという立場とパラメタ数を増やすことで単調関係を満たそうとする立場とがある。前者の立場に立った効用関数推定法としても色々な方法が提案されているが、ここでは新しいパラメタ決定尺度を提案し、Kruskalの単調回帰原理を用いる方法 [2] とJohnsonの対の順序関係保存度に基づく方法 [4] と比較する。後者の立場としてはパラメタ自身があいまいさを持つとしてパラメタをファジィ数とする方法を提案する。

コンジョイント分析法は各個人の持つ評価構造を解明する手法として発展してきたため、個々の回答に適合する指標については文献[2],[4]以外にも多くの検討が行われてきたが、回答全体を扱える指標については検討されていない。そこで回答全体を扱える総合評価指標についても提案する。

2. モデルの概要

選択対象 i の全体効用を U_i とし，j 番目の属性に対する部分効用を u_{ij} とすると，

$$U_i = F(u_{i1}, u_{i2}, \ldots, u_{ir}) \tag{1}$$

と表現できる（属性数は r とする）が，これまでの検討結果に合わせて加法的結合ルール

$$U_i = \sum_{j=1}^{r} u_{ij} \tag{2}$$

に従う場合を論じる．選択対象 i の選好度を P_i，選択順序を S_i とすると，

$$P_i > P_k \quad \text{または} \quad S_k > S_i \quad \text{ならば} \quad U_i \geqq U_k \tag{3}$$

となるように部分効用 u_{ij} を決める．

選択対象 i の選好度はその属性で決定されるものとする。N 個の属性は質的変数（カテゴリカル・データ）で表現され，M 個の属性は量的変数で表現される場合を考える．属性 j は n_j 個のカテゴリを持つものとする．選択対象 i の全体効用は

$$U_i = \sum_{j=1}^{N}\sum_{k=1}^{n_j} a_{jk}\delta_{jk}(i) + \sum_{h=1}^{M} b_h x_h(i) \tag{4}$$

で与えられるものとすると

$$1 \leqq j \leqq N \quad \text{では} \quad u_{ij} = \sum_{k=1}^{n_j} a_{jk}\delta_{jk}(i) \tag{5}$$

$$N+1 \leqq j \leqq N+M \quad \text{では} \quad u_{ij} = b_{j-N}x_{j-N}(i) \tag{6}$$

である．ただし，

$\delta_{jk}(i) = 1$: 対象 i は属性 j の分類 k に属する

0 : その他

$x_h(i)$: 対象 i の属性 h の値

である．

3．ボイスメール・サービス選好度の分析

制約条件付きでコンジョイント分析を行う場合に対の順序関係保存度に基づく方法 (TRADE-OFF)[4] では対象ごとの効用に差が出ない事例があった．そこで TRADE-OFF の指標

$$\theta^2 = \sum_{i<j} E_{ij}(U_i - U_j)^2 / \sum_{i<j}(U_i - U_j)^2 \tag{7}$$

ただし、$E_{ij} = 1 : (P_i - P_j)(U_i - U_j) < 0$ (8)

0 : その他

を若干、変更した指標を提案し、ボイスメール・サービスを例にとってその指標の効果を検討する．

3.1 新しい指標の提案

対の順序関係保存度に基づく方法 (TRADE-OFF) では対象 i の効用 U_i を式 (4) とするとき、E_{ij} を使った式 (7) の指標が基準値以下になるように a_{jk}, b_h を求める．ここで問題となるのはほとんどの U_k に同じ値を与えたとき θ^2 は小さくなることがあり得るということである．そこで

$$\xi^2 = \sum_{i<j} F_{ij}(U_iP_i - U_jP_j)^2 / \sum_{i<j}(U_iP_i - U_jP_j)^2 \tag{9}$$

$F_{ij} = 1 : (P_i - P_j)(U_i - U_j) \leq 0$

0 : その他 (10)

なる規準を使うと、$P_i \neq P_j$ のときには $U_i = U_j$ とおくことは分子の量をふやすことになる．もちろん

$$(P_i - P_j)(U_i - U_j) > 0$$

であれば $\{F_{ij}=0\}$ なので分子の量はふえない．また $F_{ij}=1$ の場合に式(9)の分子の $(U_iP_i-U_jP_j)^2$ を零に近付けることは ξ^2 を小さくする方向だが、これは U と P の関係からは望ましくないので

$$\omega^2=\sum_{i<j}F_{ij}(U_iP_j-U_jP_i)^2\Big/\sum_{i<j}\{F_{ij}(U_iP_j-U_jP_i)^2+(1-F_{ij})(U_iP_i-U_jP_j)^2\} \qquad (11)$$

を提案する．ところで、E_{ij} は Y に依存しない不連続な量として扱っており、F_{ij} でもそのようにすると、指標のパラメタによる偏微分を考えるときに不連続点での値が問題となる。これを改善するため、十分大きな L（例えば L=10^5）をとって

$$F_{ij}=[1+\exp\{L(P_i-P_j)(U_i-U_j)\}]^{-1} \qquad (12)$$

と近似することが考えられる。ここでは ω^2 の計算にこの考え方を取り入れた。

表1　8種類のボイスメール・サービス(A-H)と対応する属性

j	I (入会金)	II (基本料)	III (使用料／分)	IV (message 数)	V (保存期間)
A	2万円	5千円	50円	50	1箇月
B	2万円	5千円	50円	5	1週間
C	2万円	2百円	5円	5	1箇月
D	2万円	2百円	5円	50	1週間
E	2千円	5千円	5円	5	1週間
F	2千円	5千円	5円	50	1箇月
G	2千円	2百円	50円	50	1週間
H	2千円	2百円	50円	5	1箇月

3.2　分析結果

ボイスメール・サービスを表1に示す5属性で規定する．以下ではどの属性も質的変数（カテゴリカルデータ）として取り扱う。5属性を組み合わせた8種類のサービスに対する選好度 P=(P1,P2,…,P8) データを15企業から得た．

属性の定性的優位性から

$$a_{11}\geqq a_{12},\ a_{21}\geqq a_{22},\ a_{31}\geqq a_{32},\ a_{41}\leqq a_{42},\ a_{51}\leqq a_{52} \qquad (13)$$

の条件下で θ^2 の最小化を行うと企業7のデータ

(P=(5,3,1,6,4,7,8,2))

に対しては　$a_{42}=1$、その他の $a_{jk}=0$

となり、　$U_A=U_D=U_F=U_G=1$；その他の $U_i=0$

なので　$\theta^2=0$

となる．従って、指標 θ^2 による P の推定値 P_θ は

P_θ=(5,1,1,5,1,5,5,1)

となる。これに対して指標ξ^2を採用すると

$a_{11}=0.1203,\ a_{42}=0.9927$, その他の $a_{jk}=0$

となり、ξ^2によるPの推定値P_ξは

$P_\xi=(5,1,1,5,3,7,7,3)$

となるので、Pの近似としてP_ξの方がP_θよりも望ましい。この例では指標ω^2によるPの推定値P_ωはP_ξに等しい。

今回提示したサービスの効用を定性的に順序付けると

(i) サービスBは最下位：$\min_i U_i=U_B$

(ii) サービスFはサービスA、Eよりも上位：$U_F>U_A, U_E$

(iii) $U_D>U_C \longleftrightarrow U_G>U_H$; $U_D<U_C \longleftrightarrow U_G<U_H$

である．企業9の回答Pは(i),(iii)を満足せず、得られた推定値$P_\times$はそれを調整する形になっている．企業10, 11, 12, 13の回答Pは(i)を満足せず、推定値P_ξ、P_ωはそれを調整している．企業15の回答Pは(iii)を満足せず、推定値$P_\times$はそれを調整している．企業14の回答Pは(ii)を満足せず、推定値$P_\times$はそれを調整している．この外、企業5については、P_η、P_θ、P_ωは同じだが、P_ξは若干異なる。P_ξはAとHが逆転しているが、C, D, Fの順位付けは他よりも優っており、各指標の優劣は付けられない。企業8については、AとDの逆転がないということでP_ωが若干優れている。

逆に、企業13については、FとGとが逆転しているということでP_ωが若干劣っている。企業1, 2, 3, 4, 6, 7, 9, 15については指標間に差はない。

以上から、P_θ、P_ξ、P_ωとはほぼ同じで、企業11の場合にも指標ξ^2、ω^2は対応し得ることが分かる．P_ηはP_θの欠点をさらに強調したものになっており、用いないほうがよいと思われる。しかしξ^2、ω^2による場合はθ^2による場合よりも収束速度が遅いという欠点がある．

4．総合評価指標の提案

これまでは各企業や個人の選好度に着目してきたが、全回答を通して総合的に議論できるモデルを提案し、ボイスメール・サービスに適用してみる。

式(4)で示される効用における係数a_{jk}、b_hの値は回答者ごとに異なっており、回答者pを陽に表現すると

$$U_i^{(p)}=\sum_{j=1}^{N}\sum_{k=1}^{n_j} a_{jk}^{(p)}\delta_{jk}(i)+\sum_{h=1}^{M} b_h^{(p)}x_h(i) \qquad (14)$$

となる。ここでは、全回答を通して係数a_{jk}, b_hは同じ値を取るが、回答者ごとに属性に感じる価値が異なるものとし、回答者pが選択対象iに感じる効用を

$$U_i^{(p)}=\sum_{j=1}^{N} v_{pj}\sum_{k=1}^{n_j} a_{jk}\delta_{jk}(i)+\sum_{h=1}^{M} v_{ph}b_h x_h(i) \qquad (15)$$

と表現する。回答者pの選択対象iに与えた選好度を$P_i(p)$、i-j間の不一致度指標を

$F_{li}(p)$（式(10)）とし、式(11)から類推される回答者全体としての不適合度指標を

$$\omega_T{}^2 = \sum_p \sum_{i<j} F_{ij}{}^{(p)}(U_i{}^{(p)}P_j{}^{(p)} - U_j{}^{(p)}P_i{}^{(p)})^2 / \Omega \tag{16}$$

$$\Omega = \sum_{i<j} \{F_{ij}{}^{(p)}(U_i{}^{(p)}P_j{}^{(p)} - U_j{}^{(p)}P_i{}^{(p)})^2 + (1 - F_{ij}{}^{(p)})\,(U_i{}^{(p)}P_i{}^{(p)} - U_j{}^{(p)}P_j{}^{(p)})^2\} \tag{17}$$

とする。ここで回答者ごとに効用の総和は一定として

$$\sum_{i=1}^{m} U_i{}^{(p)} = m(m+1)/2 \tag{18}$$

の制約を課す。効用の自乗和を一定とすることも考えられる。

$\omega_T{}^2$ は次の手順で最小化される。

(i)$\sum_p P_i{}^{(p)}$ を被説明変数とし、$\delta_{jk}(i), x_h(i)$ を説明変数として重回帰分析（質的変数、量的変数混在時の数量化理論１類）により偏回帰係数 a_{ik}, b_h を求める。ここで、a_{jk} については $\{\min_k a_{jk}=0\}$ となるように調整される。また、$x_h(i)$ については最悪の値が零となるように調整される。

(ii) $P_i{}^{(p)}$ を被説明変数とし、$\delta_{jk}(i), x_h(i)$ を説明変数として重回帰分析（質的変数、量的変数混在時の数量化理論１類）により偏回帰係数 $a_{jk}{}^{(p)}$、$b_h{}^{(p)}$ を求める。ここで、$a_{jk}{}^{(p)}$ についても $\{\min_k a_{jk}{}^{(p)}=0\}$ となるように調整される。計算時間にこだわらなければ回答者ごとに節４の ω^2 を最小にするパラメタを求めて $a_{jk}{}^{(p)}$、$b_h{}^{(p)}$）としてもよい。

(iii) $v_{pj} = \max_k a_{jk}{}^{(p)} / \max_k a_{jk}$ 、$v_{ph} = b_h{}^{(p)} / b_h$ とする。

(iv) 式(18)を満たすように v_{pj} を変更する。

(v) a_{jk}, b_h を固定したまま、$\omega_T{}^2$ を最小にする v_{pj} を求める。

式(18)の制約を考慮しないときには

(vi) v_{pj} を固定したまま、$\omega_T{}^2$ を最小にする a_{jk}, b_h を求める

という手順を追加することが可能であるが、式(18)の制約数は一般に変数の数よりも多いため手順(vi)は採れない。即ち、a_{jk}, b_h は手順(i)で求められたものを改良できないので手順(i)の改良で対処する必要がある。

上記手順(i)～(v)をボイスメール・サービスに適用した結果、

a_{11}=0.384, a_{21}=0.202, a_{31}=0.486, a_{42}=0.759, a_{52}=0.010, その他の a_{jk}=0

となり、別途、検討したパラメタのレンジが第一位の回数とほぼ同じ傾向の値となっている。なお、$\omega_T{}^2$=0.0017 であった。企業 L の選好度 P と、その $U_i(L)$ による推定値 P_U とは節３で得られた結果とほぼ同じであり、個々の選好度を適切に推定しつつ総合指標 $\omega_T{}^2$ の最小化も達成されていると言える。なお厳密には $\omega_T{}^2$ の局所最小化しか保証されておらず、手順(iii)で $\{v_{pj}=1\}$ とした場合には若干、異なる値に収束した。データ通信サービスについても同様の結果が得られた。

これらのことから総合指標は有効であると言え、全体としての選好構造を議論でき

るようになった。

5. ファジィ・コンジョイント分析法

これまでのコンジョイント分析法では，効用を推定できなかった回答は無視されていた．たとえばMONANOVAではストレスと呼ばれる非適合度指標η，TRADE-OFFでは非適合度指標θ^2が十分小さくならないデータは無視される．LINMAPでは線形計画法が解を持たないことによりデータは無視される．

しかし，データを捨てていると，有効データは限られてしまうし，質の悪いデータのなかにも情報が含まれている．例えば，複数サービスの同順位を認めていない場合にある属性の有無だけで判断している回答者がいるとすると，その属性を持っているサービス間，あるいは持っていないサービス間の順位は無意味であるが，回答の合理性をチェックしている人から見ると別の属性の良否から来る明らかな順序付けがなされていないと見るかもしれない．そこで，積極的に同順位は認めつつ異順位を与えることが可能な場合には異順位の方が採用される方法を検討する．

これを実現するためには既存の方法ではパラメタに制約を付けることが考えられる．しかし，目的関数が非線形の場合にはそのプログラミングは容易ではないし，初期値の設定やパラメタの修正幅などには工夫がいる．そこでそれらの心配の必要のない線形計画法での取扱いが可能となるようファジィ・コンジョイント分析法を提案し，ボイスメールサービスを対象に既存手法との比較を行う．

5.1 ファジィ・コンジョイント・モデル

MONANOVA, TRADE-OFFでは，ともに目的関数が非線形のため，そのプログラミングは容易ではない．そこで線形計画法の利用を図れるようなアルゴリズムについて考える．LINMAP[5]は線形計画法を利用するアルゴリズムとして知られているが，順序逆転を表す非負の変数の和Sを最小とするアルゴリズムであるため，$\{S=0\}$ の場合にしか合理的な解は得られないので，ここでは比較対象とはしない．

ここではカテゴリカル・データだけを取り上げたが，式(4)のように数値データが含まれても同様の議論が可能である．

対象iが選好順位の順に並べられているとすると，全体効用に関しては

$$U_i > U_h \ (i < h)$$

であってほしい．U_iとU_hの順位が逆転しているとき，それはa_{jk}のあいまいさに起因していると考える．すなわち，a_{jk}は確定値（通常数）ではなく，下限値を$(\alpha_{jk} - c_{jk})$，上限値を$(\alpha_{jk} + d_{jk})$，モードをα_{jk}とする三角型ファジィ数$(\alpha_{jk} - c_{jk},\ \alpha_{jk},\ \alpha_{jk} + d_{jk})$であると考える．このとき，隣り合う対象の効用差$V_i = U_i - U_{i+1}$も三角型ファジィ数$(v_{i1}, v_{i2}, v_{i3})$で与えられる．

$V_i > 0$を，ファジィ数の場合にどのように捕らえるかが問題である．そこで次の2種

類の考え方を取り上げる．

【考え方１】三角型ファジィ数$\tilde{A}=(a_1,a_2,a_3)$の順序づけとして値 0 に関するremoval ([7], p.36)

$$R(\tilde{A},0)=(a_1+2a_2+a_3)/4 \tag{19}$$

を用いる．すなわち

$R(\tilde{U}_i,0)>R(\tilde{U}_{i+1},0)$ ならば $\tilde{U}_i>\tilde{U}_{i+1}$

と考える．

このとき，つぎのようなLP問題が考えられる．

《FR1》

[目的] $$\max\{\sum_{i=1}^{n-1}s_i-\sum_{j,k}(c_{jk}+d_{jk})\} \tag{20}$$

[制約] $$R(\tilde{U}_i,0)-R(\tilde{U}_{i+1},0)-s_i=\varepsilon \tag{21}$$

$$R(\tilde{U}_1,0)=n-1\ ;\ R(\tilde{U}_n,0)=0\ ;\ \alpha_{jk},c_{jk},d_{jk}\geq 0 \tag{22}$$

【考え方２】v_{i1}のiに関する最小値pをできるだけ大きくし，かつ曖昧さ $\sum_{j,k}(c_{jk}+d_{jk})$ をできるだけ小さくしたい．すなわち，目的はmax pかつ，min $\sum_{j,k}(c_{jk}+d_{jk})$である．

この２目的関数はスラック変数s_iも考慮して統合でき，次のLP問題になる．

《FL1》

[目的] $$\max\{p-\sum_i s_i-\sum_{j,k}(c_{jk}+d_{jk})\} \tag{23}$$

[制約] $$v_{i1}-s_i=p\ ;s_i\geq 0 \tag{24}$$

$$R(\tilde{U}_1,0)=n-1\ ;\ R(\tilde{U}_n,0)=0\ ;\alpha_{jk},c_{jk},d_{jk}\geq 0 \tag{25}$$

ここで，s_iとpの役割を分離するために，式(23)における $\sum_i s_i$ の係数が負であることに注意しなければならない．

また，なるべく選好順位が隣合うサービスの効用差が均等になることを狙った以下のLP問題も考えられる．

《FR2》

[目的] $$\max\{p-\sum_i s_i-\sum_{j,k}(c_{jk}+d_{jk})-\varepsilon_u-\varepsilon_l\} \tag{26}$$

[制約] 式(21), (22)のほかに

$$1-\varepsilon_l\leq R(\tilde{U}_i,0)-R(\tilde{U}_{i+1},0)\leq 1+\varepsilon_u \tag{27}$$

《FL2》

[目的] $$\max\{p-\sum_i s_i-\sum_{j,k}(c_{jk}+d_{jk})-\varepsilon_u-\varepsilon_l\} \tag{28}$$

[制約] 式(24), (25)のほかに

$$1-\varepsilon_l \leq v_{i2} \leq 1+\varepsilon_u ; \varepsilon_u, \varepsilon_l \geq 0 \tag{29}$$

式(22), (25)はパラメタの大きさを揃えるために採用した制約であるが，これについてはremoval ではなくモードを用いたり，パラメタの和を一定にするなど別の制約にしてもよい．しかし，別の尺度を用いると結果には若干の差が出ることがある．

5.2 ボイスメールサービスへの適用

節4で論じた 8 種類のボイスメールサービスに対してコンジョイント分析を行う．
この場合のパラメタは　$Y=(a_{11}, a_{12}, a_{21}, a_{22}, a_{31}, a_{32}, a_{41}, a_{42}, a_{51}, a_{52})$
であり，属性の優劣から　$a_{12}, a_{22}, a_{32}, a_{42}, a_{52}=0 ; a_{11}, a_{21}, a_{31}, a_{41}, a_{51} \geqq 0$
の制約を付ける．

結果を分析した結果、下記のことが分かった．

(1) 企業1～4 についてはどの方法でも順位が再現されている．

(2) 企業5～8 についてはFR1を除き，矛盾のない全体効用値が得られている．FR1では，同順位が多く，式(21)の ε に値0.01を設定してもしなくても大して結果には差がなかった．

(3) 企業9 の回答は条件(i),(iii)を満足せず，評価結果は上記(2)と同様である．

(4) 企業10，11，12，13 の回答は条件 (i)を満足せず，既存手法では積極的に同順位を認めるMONANOVA以外では全体効用値と回答順位とで逆転しているところが見受けられる．特にU_{NLS}は逆転の度合いが大きい．ファジィ・コンジョイント・モデルではFR1で{ε=0.01}と無理やり全体効用に差を付けようとすると解が求められなかった．

(5) 企業14 の回答は(ii)を満足せず，評価結果は上記(4)と同様である．

(6) 企業15 の回答は(iii)を満足していないが，FR1を除き矛盾のない全体効用値が得られている．特に{ ε =0.01}では解が求められなかった．

6．むすび

コンジョイント分析法をボイスメール・サービス選好度の分析に応用した。その際に既存のコンジョイント分析法の欠点を改善する新しい尺度および総合評価指標の提案を行った。総合評価指標についてはほかの指標やパラメタの導出法も考えられ、今後の検討が必要である。用いたデータは全対象を順序づけた「フルプロファイル法」[6]と呼ばれる方法によっているが、今後はコンピュータ・インタビューに馴染んだ「一対比較法」に基づくデータ収集法も考慮する必要がある。

定性的整合性を満たさない回答についても無視せずに全体効用値を与えられる方法も検討した．その結果，既存手法では積極的に同順位を認める制約付きのMONANOVA，ファジィ・コンジョイント・モデルではFR2, FL1, FL2 が矛盾のない

全体効用値を与えてくれることが分かった．ファジィ・コンジョイント・モデルは線形計画法のプログラムを持っていれば容易に計算できるので制約付きのMONANOVAよりも推奨したい．数少ない比較なので断定はできないが，ファジィ・コンジョインの中でも優劣を付ければ全体効用値の同順位が少ないという意味でFR2が最も優れていた．

ここで示したファジィ・コンジョイント・モデルをさらに多くのデータに適用してみた結果、まだ、説明できない回答もあった。そのようなもののいくつかは更にパラメタを増やして特定の属性を持つ場合にはプレミアムを付けると言った方法により説明ができた。例えば最も低料金のサービスには通常の部分効用よりも高めの数値を与えるなどである。もちろん、パラメタをどんどん増やしていけばすべての回答を説明できるモデルはできあがるが、そのようなモデルと他の簡単なモデルとを総合化する場合に問題があり、やたらにパラメタは増やすべきでないであろう。

文　献

[1] Luce, R.D. and Tukey, J.W. : "Simultaneous conjoint measurement : A new type of fundamental measurement", J. Math.Psychology, 1, pp.1-27 (1964).

[2] Kruskal, J.B. : "Analysis of factorial experiments by estimating monotone transformations of the data", *J. Royal Statistical Society*, Series B-27, pp.251-263 (1965).

[3] Kruskal, J.B. and F.J. Carmone, Jr. MONANOVA : A FORTRAN IV Program for Monotone Analysis of Variance (Non-Metric Analysis of Factorial Experiments), *Behavioral Science*, Vol.14, (1969), 165-166.

[4] Johnson, R.M. : "A simple method for pairwise monotone regression", *Psychometrika*, 40, 2, pp.163-8 (1975).

[5] Srinivasan, V. and Shocker, A.D. : "Linear programming techniques for multidimensional analysis of preferences", *Psychometrika*, 38, 3, pp.337-69 (1973).

[6] 小川孔輔，服部正太，山中正彦 ："CJMにおける２つのデータ収集法の比較"，マーケティング・サイエンス，38, pp.51-59 (1991).

[7] Kaufmann, A. and Gupta, M.M.著，田中・松岡 訳：ファジィ数理と応用，オーム社 (1992).